Regents Junior Faculty Fellowship

Ryan Giordano

I propose to spend the summer of 2024 working on two collaborative research projects. The first, “neural network classifiers for Bayesian posteriors,” promises to introduce a completely new set of Bayesian inference techniques with different computational tradeoffs than existing methods. The second, “black–box computable diagnostic weights for survey sampling,” will bring a much–needed set of diagnostic tools to the vast majority of modern applied survey sampling. These two projects are different in scope — the first represents ground–breaking methodological research, and the second an application of my existing research to an urgent applied problem — but each rests on and contributes to my existing work on approximate Bayesian computation and sensitivity analysis.

Neural network classifiers for Bayesian posteriors

Bayesian statistical techniques are a conceptually powerful set of tools for representing and quantifying uncertainty, and are increasingly popular across the physical and social sciences (CITE). Often, a statistical analysis involves a single quantity of interest, such as the effect of a policy intervention (CITE), the type of an astronomical object (CITE), the outcome of an election (CITE), or the identity of an ancestral genetic population (CITE), and Bayesian statistics abile to propagate uncertainty from a any unknown latent modeling quantities to the final estimate. But this conceptual strength is a computational weakness, since even approximately accounting for a large number of latent quantities is computationally intensive. Bayesian estimates often take hours to days to compute, and it is of considerable interest to develop computationally efficient, approximate Bayesian procedures.

I have recently shown that a slight modification of the preceding procedure can use neural network (NN) classifiers to learn low-dimensional marginals of Bayesian posteriors using only simulated data. The idea is based on the estimation of log–likelihood ratios in simulation–based inference. Furthermore, the accuracy of the approximateion can be easily checked using simulation–based calibration (SBC), a well–known Bayesian validation procedure. Note that SBC is rarely used in practice, since it is prohibitively computationally expensive in most classical Bayesian procedures. However, for a posterior approximation based on a NN classifier, SBC is computationally cheap. For the cost of training a NN classifier, one can get closed-form estimates of Bayesian posterior marginals with strong, computable statistical accuracy guarantees.

To my knowledge, there are no existing techniques that offer the advantages of my proposed method. There has been recent interest in statistical procedures that improve computability by accounting only approximately for the dependence between large numbers of latent variables. However, these approximations come without accuracy guarantees, and are known to be inaccurate in certain practical cases. The vast majority of existing research in approximate Bayesian computation attempts to model the entire high–dimensional distribution, even when only one variable is of interest.

Though superficially distinct, recent work in simulation–based inference points towards a completely new set of Bayesian inference techniques for low–dimensional marginals. Simulation–based inference is developed for problems which can be simulated, but for which no likelihood is available. By a clever construction of simualted data, researchers are able to estimate log likelihood ratios using a neural network classifier (CITE). These likelihood ratios are then used to compute maximum likelihood (points) estimates, rather than full distributions, and it is difficult to assess whether the likelihood ratios are accurate.

Note that Bayesian approaches to simulation–based inference are not new, but existing techniques are built on high–dimensional density approximation, such as normalizing flows. As with other approximate inference techniques, this set of tools approximates the entire posterior, even when only a low–dimensional marginal is of interest. To the best of my and my collaborator's knowledge, the technique above is new.

A large number of applications are immediate candidates for the above ideas.

Black–box computable diagnostic weights for survey sampling

Most modern surveys — such as polling about the upcoming presidental election — must overcome the fact that their sampled population is different from the target population. For example, the set of people responding to an internet survey about political preferences is likely to differ systematically from the full population of voters. As a consequence, survey responses are typically reweighted by key demographic variables in order to become more representative. As a diagnostic, it is very useful to be able to check that these weights are able to "balance" demographic quantities. Unfortunately, the best statistical procedures for inferring the polling responses of rare demographic groups are nonlinear, and so do not readily admit diagnostic weights.

References